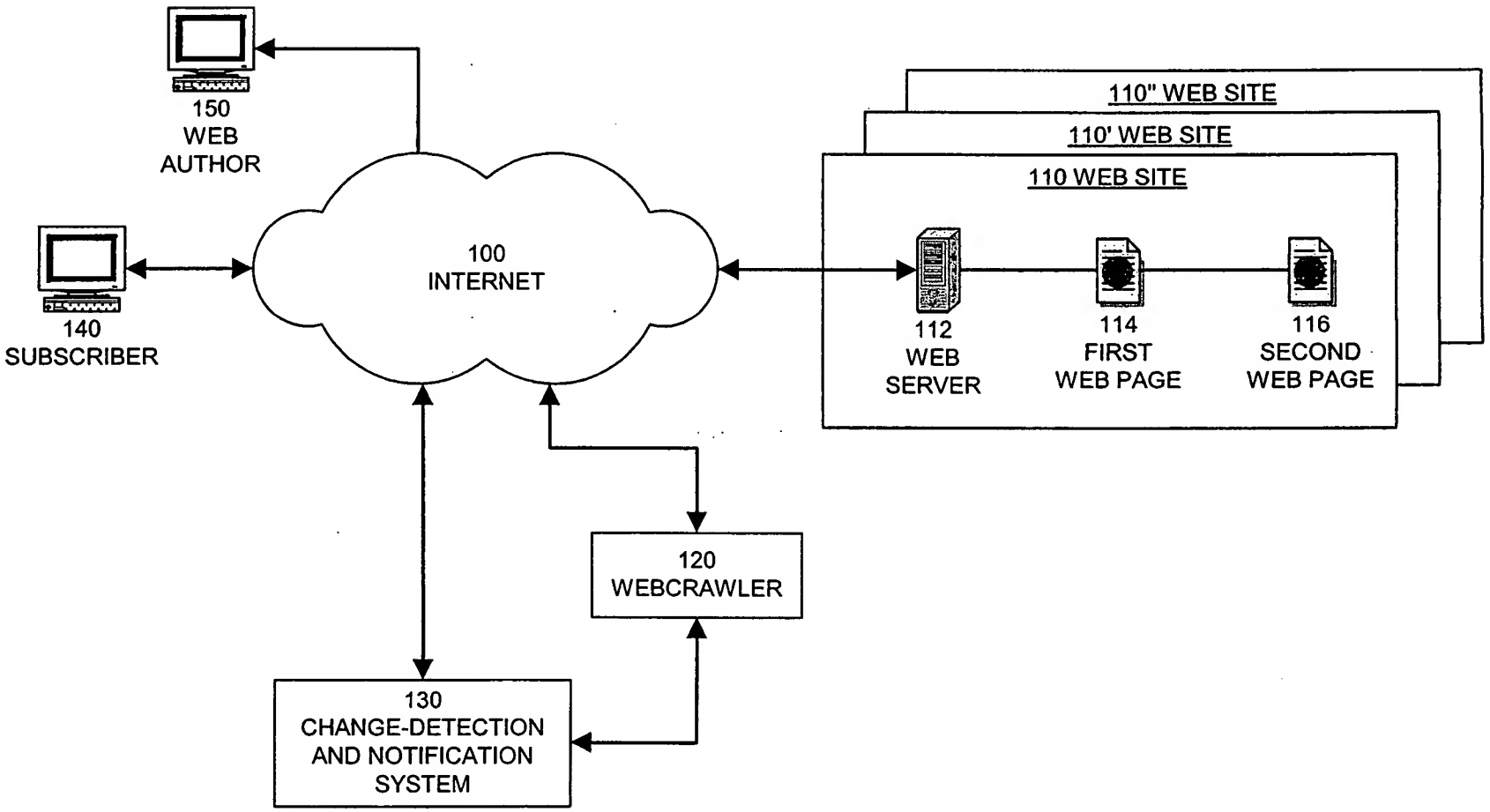
150
WEB
AUTHOR
140
SUBSCRIBER
100
INTERNET
110" WEB SITE
110' WEB SITE
110 WEB SITE
112
WEB
SERVER
114
FIRST
WEB PAGE
116
SECOND
WEB PAGE
120
WEBCRAWLER
130
CHANGE-DETECTION
AND NOTIFICATION
SYSTEM

FIG. 1

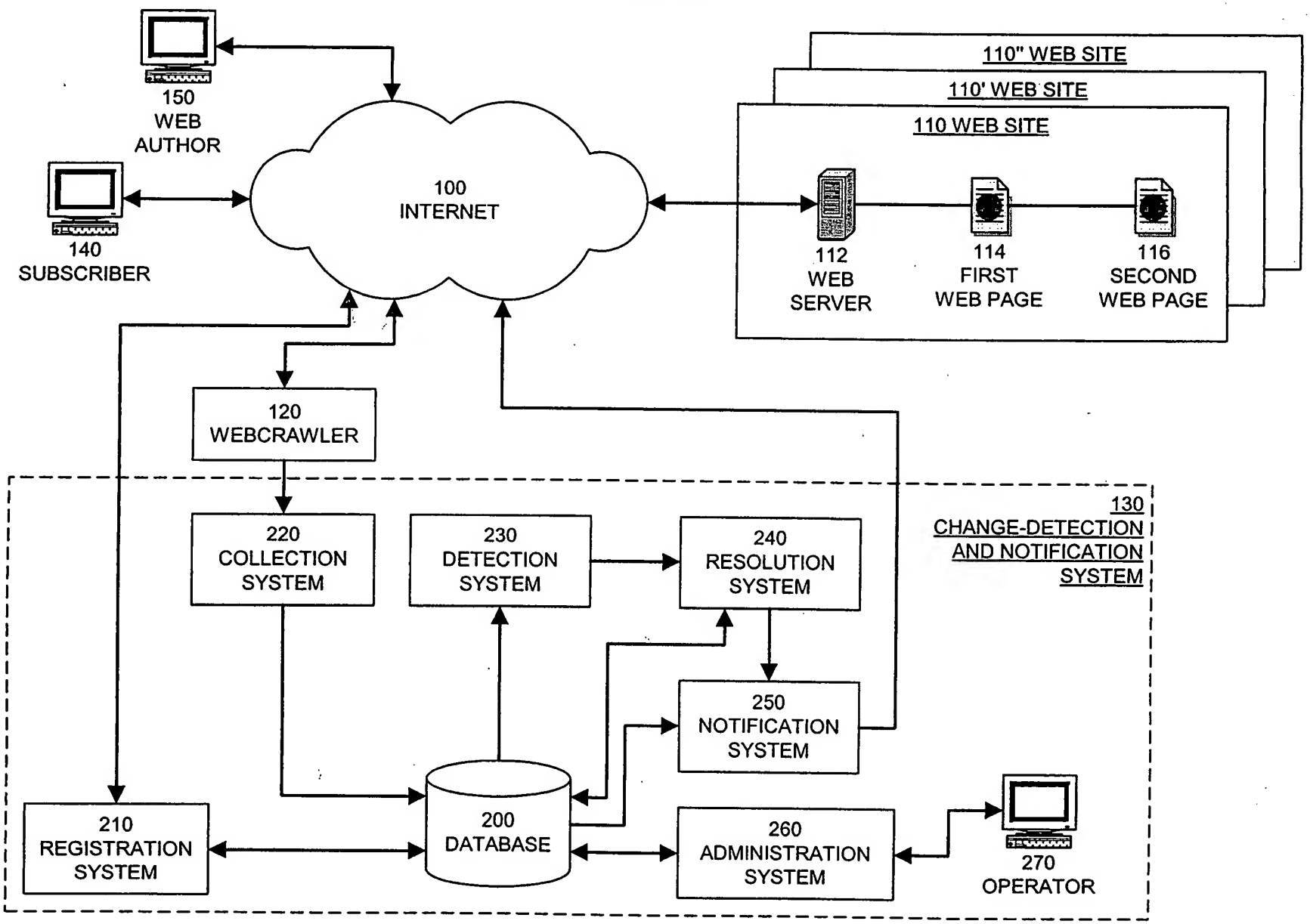
150
WEB
AUTHOR
140
SUBSCRIBER
100
INTERNET
110" WEB SITE
110' WEB SITE
110 WEB SITE
112
WEB
SERVER
114
FIRST
WEB PAGE
116
SECOND
WEB PAGE
120
WEBCRAWLER
130
CHANGE-DETECTION
AND NOTIFICATION
SYSTEM
220
COLLECTION
SYSTEM
230
DETECTION
SYSTEM
240
RESOLUTION
SYSTEM
250
NOTIFICATION
SYSTEM
210
REGISTRATION
SYSTEM
200
DATABASE
260
ADMINISTRATION
SYSTEM
270
OPERATOR

FIG. 2

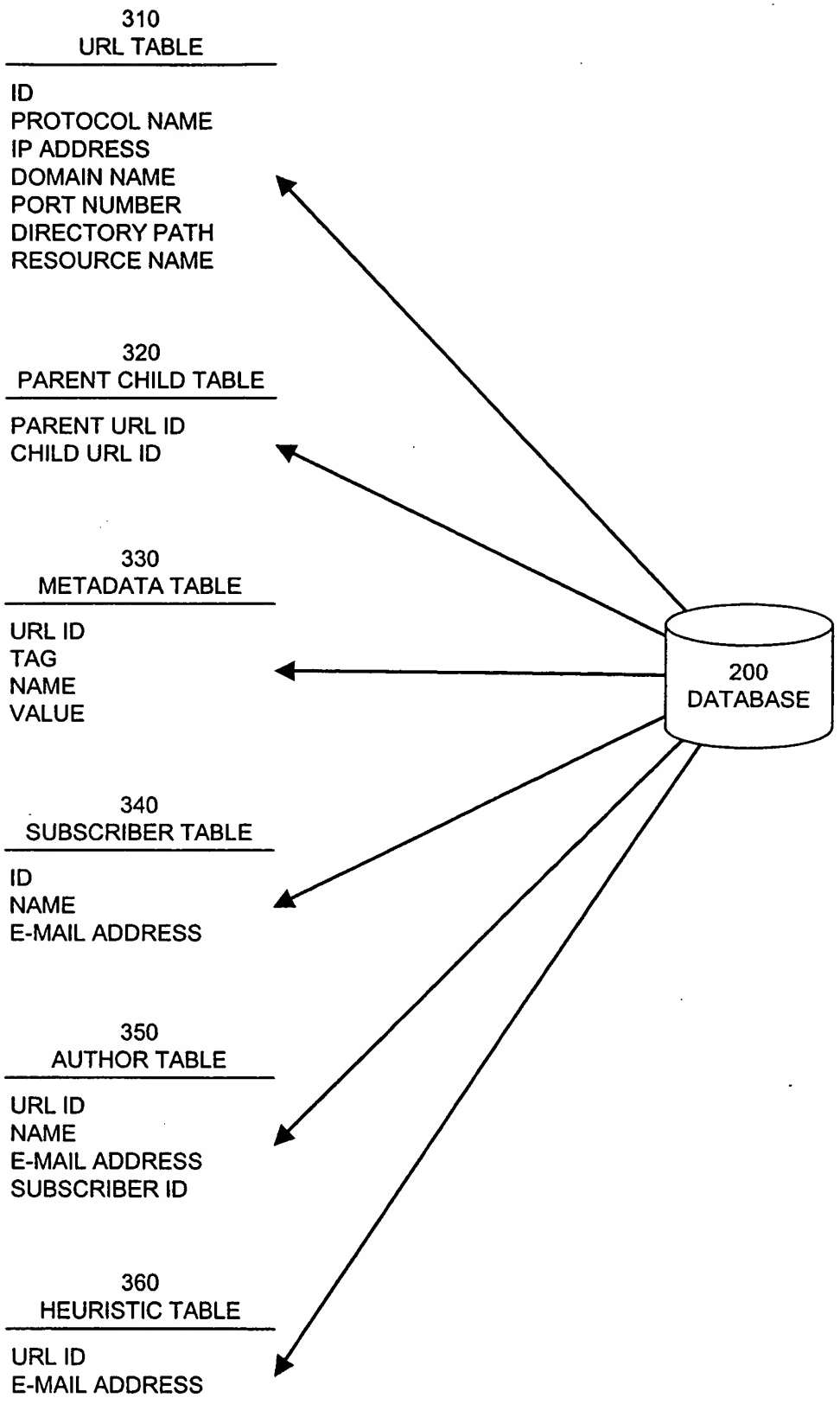
310
URL TABLE
ID
PROTOCOL NAME
IP ADDRESS
DOMAIN NAME
PORT NUMBER
DIRECTORY PATH
RESOURCE NAME
320
PARENT CHILD TABLE
PARENT URL ID
CHILD URL ID
330
METADATA TABLE
URL ID
TAG
NAME
VALUE
340
SUBSCRIBER TABLE
ID
NAME
E-MAIL ADDRESS
350
AUTHOR TABLE
URL ID
NAME
E-MAIL ADDRESS
SUBSCRIBER ID
360
HEURISTIC TABLE
URL ID
E-MAIL ADDRESS
200
DATABASE

FIG. 3

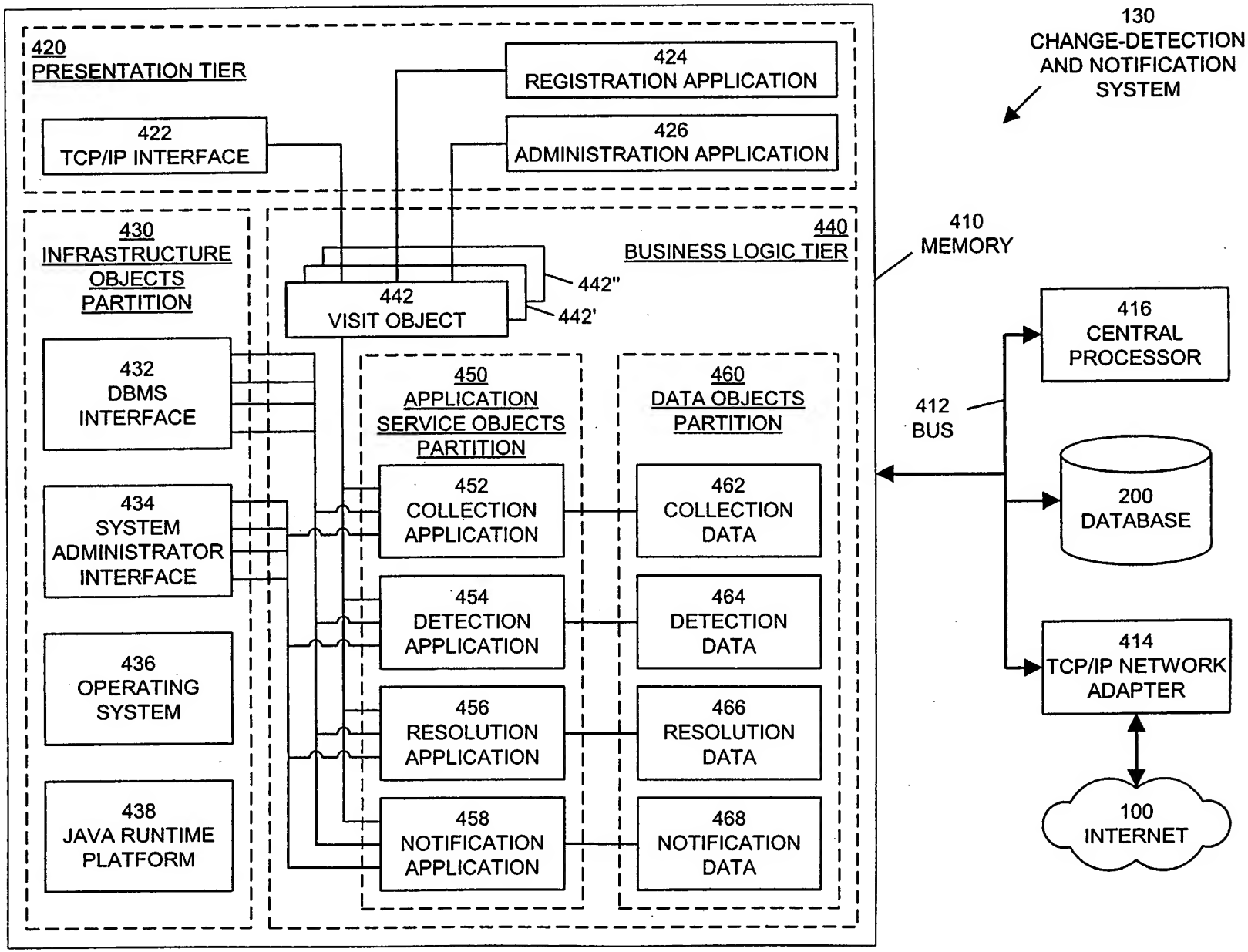
130
CHANGE-DETECTION AND NOTIFICATION SYSTEM
420
PRESENTATION TIER
424
REGISTRATION APPLICATION
422
TCP/IP INTERFACE
426
ADMINISTRATION APPLICATION
410
MEMORY
430
INFRASTRUCTURE OBJECTS PARTITION
440
BUSINESS LOGIC TIER
442"
442'
442
VISIT OBJECT
416
CENTRAL PROCESSOR
412
BUS
450
APPLICATION SERVICE OBJECTS PARTITION
460
DATA OBJECTS PARTITION
432
DBMS INTERFACE
452
COLLECTION APPLICATION
462
COLLECTION DATA
200
DATABASE
434
SYSTEM ADMINISTRATOR INTERFACE
454
DETECTION APPLICATION
464
DETECTION DATA
436
OPERATING SYSTEM
456
RESOLUTION APPLICATION
466
RESOLUTION DATA
414
TCP/IP NETWORK ADAPTER
438
JAVA RUNTIME PLATFORM
458
NOTIFICATION APPLICATION
468
NOTIFICATION DATA
100
INTERNET

FIG. 4

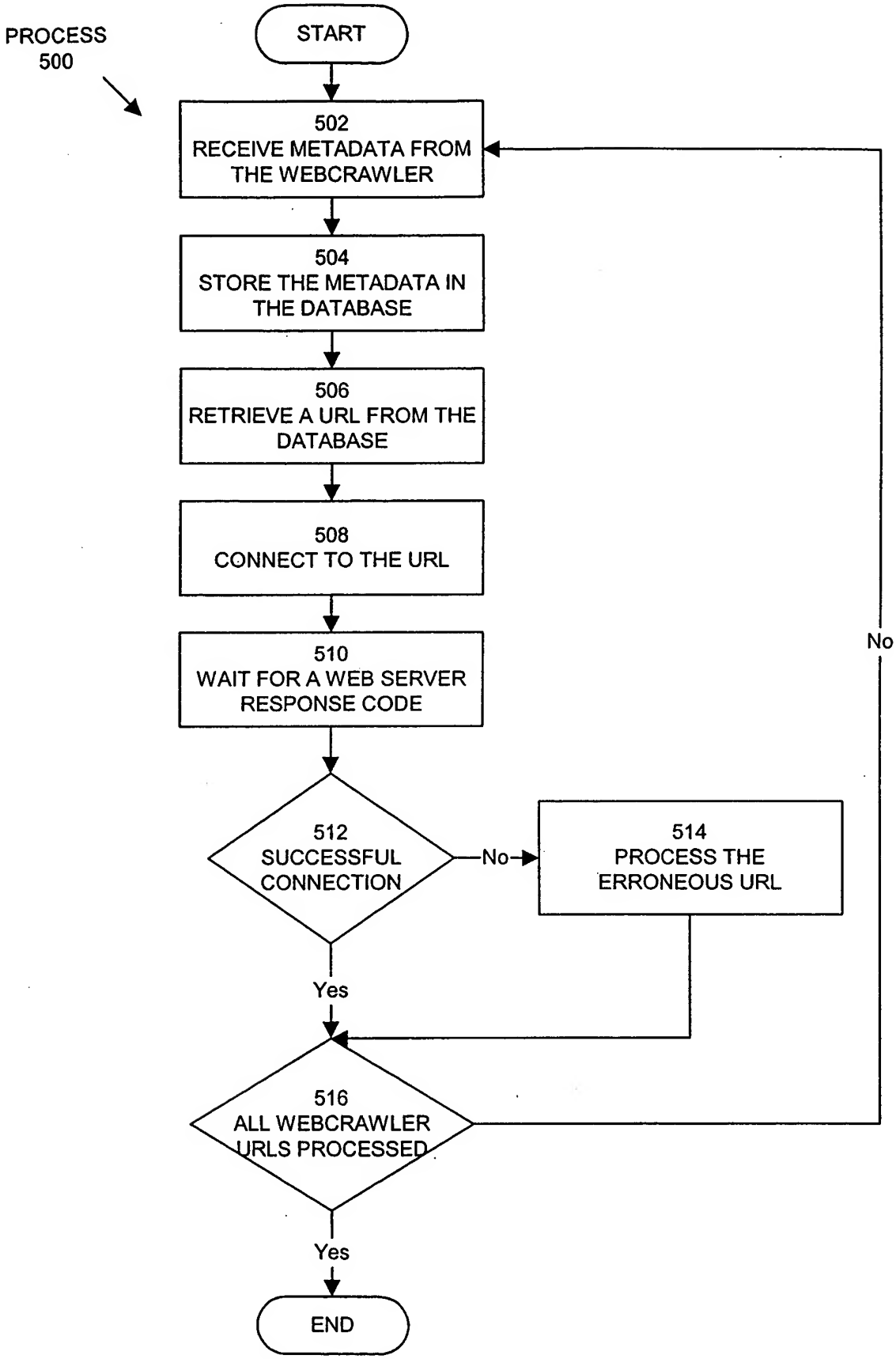
PROCESS
500
START
502
RECEIVE METADATA FROM THE WEBCRAWLER
504
STORE THE METADATA IN THE DATABASE
506
RETRIEVE A URL FROM THE DATABASE
508
CONNECT TO THE URL
510
WAIT FOR A WEB SERVER RESPONSE CODE
512
SUCCESSFUL CONNECTION
No
514
PROCESS THE ERRONEOUS URL
Yes
516
ALL WEBCRAWLER URLS PROCESSED
No
Yes
END

FIG. 5A

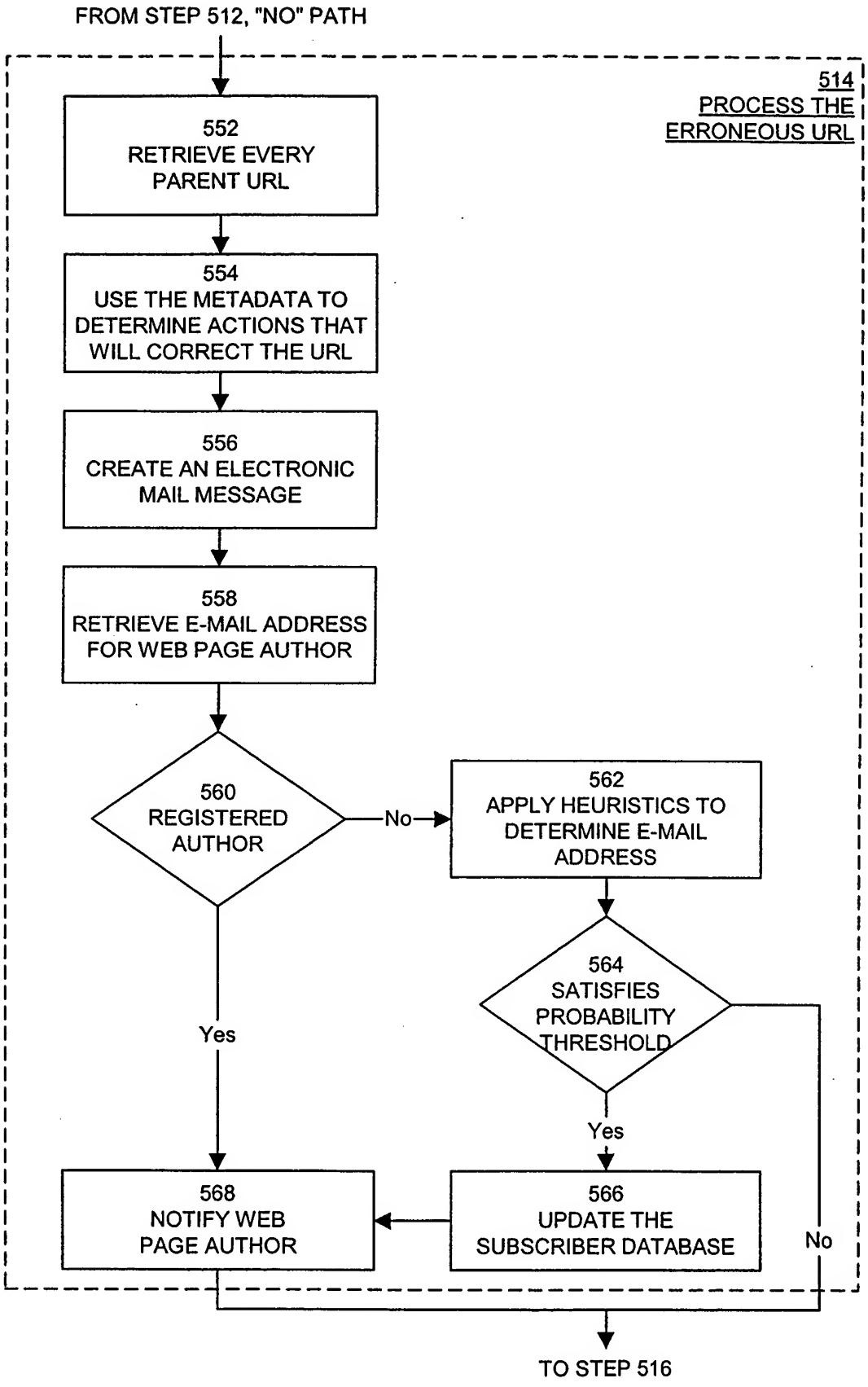
FROM STEP 512, "NO" PATH
514
PROCESS THE ERRONEOUS URL
552
RETRIEVE EVERY PARENT URL
554
USE THE METADATA TO DETERMINE ACTIONS THAT WILL CORRECT THE URL
556
CREATE AN ELECTRONIC MAIL MESSAGE
558
RETRIEVE E-MAIL ADDRESS FOR WEB PAGE AUTHOR
560
REGISTERED AUTHOR
No
562
APPLY HEURISTICS TO DETERMINE E-MAIL ADDRESS
564
SATISFIES PROBABILITY THRESHOLD
Yes
Yes
No
566
UPDATE THE SUBSCRIBER DATABASE
568
NOTIFY WEB PAGE AUTHOR
TO STEP 516

FIG. 5B